WELFARE ANALYSIS OF CHANGES IN HEALTH COINSURANCE RATES

PREPARED FOR THE OFFICE OF ECONOMIC OPPORTUNITY

KENNETH ARROW

R-1281-OEO
NOVEMBER 1973

Rand
SANTA MONICA, CA. 90406

The research reported herein was performed pursuant to a grant from the Office of Economic Opportunity, Washington, D.C. Reports of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

WELFARE ANALYSIS OF CHANGES IN HEALTH COINSURANCE RATES

PREPARED FOR THE OFFICE OF ECONOMIC OPPORTUNITY

KENNETH ARROW

R-1281-OEO
NOVEMBER 1973

Rand
SANTA MONICA, CA. 90406

PREFACE

This report studies the desirability of various provisions in health insurance policies. It was prepared for the Office of Economic Opportunity as part of a project to study the financing of medical care services for the poor and near-poor.

The study concentrates on the effects of coinsurance rates on the efficiency of allocating medical care and accompanying financial risks. Increasing the coinsurance rate has two effects: It increases the incentive for a person to economize on medical care, and it increases his financial risks. The analysis of these two effects is further complicated by the fact that an increased coinsurance rate decreases the demand for medical services and therefore lowers the price for medical services below what it would otherwise be. The lower price in turn offsets both of the two effects already noted. The effect on prices in turn depends on the extent to which the supply of medical services responds to price changes. In particular, if the supply of medical services is unresponsive to price changes, there will be neither an economizing of medical services nor an increase in financial risks but only reduction in payments to medical services.

This report extends earlier work (Kenneth J. Arrow, *Optimal Insurance and Generalized Deductibles*, R-1108-OEO, February 1973) to take account of the price effect of insurance on the demand for medical services, in conformity with the considerable evidence of the importance of this effect (see C. E. Phelps and J. P. Newhouse, *Coinsurance and the Demand for Medical Services*, R-964-OEO, April 1973; J. P. Newhouse and C. E. Phelps, *Price and Income Elasticities for Medical Care Services*, R-1197-NC, forthcoming).

Kenneth J. Arrow is Professor of Economics at Harvard University and a consultant to The Rand Corporation.

SUMMARY

This study examines the welfare implications of changes in the coinsurance rate of health insurance policies. Only efficiency aspects are studied; distributional problems are ignored.

The basic function of health insurance is the reduction of uncertainty; other things being equal, individuals prefer and are willing to pay for a reduction in their financial risks. To that extent, a reduction in the coinsurance rate would represent a welfare gain. However, given that illness has occurred, insurance constitutes a subsidy to one form of consumption and therefore implies an efficiency loss whose magnitude depends upon supply as well as demand conditions.

To get a precise expression for the net welfare change associated with a change in coinsurance rate, it is necessary to formulate the problem as a miniature general equilibrium model, with both supply and demand considerations made explicit. Account must be taken of the random factors in demand, the financing of health insurance (here assumed to be by lump-sum taxation), the elasticity of supply, and the determination of medical prices through supply and demand. For each coinsurance rate, there is an equilibrium price for medical services. Each individual, given his income net of the taxes needed to pay for health insurance, has a demand for medical services in each state of nature and therefore an expected utility, taking into account uncertainty as to health and medical costs. The study evaluates the change in expected utility as the coinsurance rate changes (see especially Theorems 1 and 3).

If the supply of medical services is totally inelastic, then a change in coinsurance rates has no efficiency effect whatsoever. The price of medical services charged by the seller changes just enough so that the coinsurance payment (the price to the buyer) remains constant; hence, there is no effect on demand or on financial risk. When supply is totally inelastic, changes in coinsurance rates affect only the distribution of income between suppliers of medical services and the rest of the population.

ACKNOWLEDGMENTS

The author is indebted to Charles Phelps and John Stein for many helpful comments, especially with regard to implicit assumptions and to exposition.

CONTENTS

I. INTRODUCTION

The following assumptions will be made: (1) The health status of an individual is a random variable whose distribution is independent of prices and income;[1] the individual aims to maximize his expected utility. (2) The utility depends upon the amount of goods other than health care, the amount of health care, and the state of health. (3) The insurance offered takes the form of reimbursing the expenditure on health care by a fixed proportion. (4) Given this coinsurance rate, the individual freely chooses the amount of medical care he wants after knowing his health status. (5) The health insurance payments are financed by lump-sum taxes.

I ignore distributional considerations and assume a single person in the economy. The interaction between distribution and insurance needs separate analysis.

This study investigates the gain or loss of welfare associated with a small change in the coinsurance rate. For this purpose, it is clearly necessary to consider both supply and demand. The rapid rate of increase of the prices of medical services since the introduction of Medicare and Medicaid can possibly be interpreted as the response of a market with relatively inelastic supply to a sudden increase of demand; if medical supply were highly elastic, the consequences for demand and therefore for efficiency could well have been very different.

I have constructed a miniature general equilibrium model of the economy disaggregated only into medical and nonmedical service markets. On the supply side, the main issue is the transformation between medical and nonmedical services. The hypothesis of perfect elasticity of supply has been implicit in most previous work. The general case is treated here. The particular case in which the elasticity of supply

[1] This assumption may be false, if higher income leads to better living conditions or to preventive medicine, which decreases the probability of serious illnesses. In the present context, the income effects are those arising from changes in the coinsurance rate and could not reasonably loom large.

is zero turns out to have a property that may be surprising at first glance, although it is not hard to see that it is true: Welfare, and indeed the allocation of resources as a whole, is totally independent of the coinsurance rate. All a change in the rate does is to transfer purchasing power between the medical and other sectors of the economy.

The welfare gains and losses have been treated in a paper by Feldstein (1973), although the theoretical basis of the calculation is not set forth too exactly. He also considers the nonoptimal behavior of nonprofit institutions. I shall analyze Feldstein's results in another paper.

II. EXPLICIT FORMULATION OF THE MODEL

Let

s = state of health,

x_{1s} = demand for goods other than health care by individual in state s,

x_{2s} = demand for medical care in state s,

x_1 = supply of goods other than medical care,

x_2 = supply of medical care,

$U(x_{1s}, x_{2s}, s)$ = utility in health state s if x_{1s}, x_{2s} is consumed in that state,

p = price of medical care received by seller,

q = price of medical care paid by buyer,

T = lump-sum taxes needed to finance health insurance,

y = income after taxes.

Prices, taxes, and income are measured with "other goods" as numeraire. To avoid distributional considerations, I assume all individuals have identical endowments and identical utility functions. I further assume a very large population, with the state of health varying independently from individual to individual. Measure all quantity variables on a per capita basis. Then the aggregate demand for commodity i is $E(x_{is})$, which is nonstochastic and is, in equilibrium, equal to the supply, x_i.

$$E(x_{is}) = x_i \ (i = 1, 2) . \tag{1}$$

I assume that the insurance policies specify

$$r = q/p = \text{coinsurance rate},$$

but the actual values of p and q are determined by market forces, to satisfy (1). The total cost of insurance (per capita) is then $(p - q)\ x_2$, or,

$$T = (1 - r)px_2 \ . \tag{2}$$

The representative individual derives his income by selling x_1 units of goods other than medical care (more exactly, the resources that will produce x_1 units of other goods) and x_2 units of medical care, the former at a price equal to 1, the latter at price p. Hence,

$$y = x_1 + px_2 - T = x_1 + rpx_2 = x_1 + qx_2 \ . \tag{3}$$

I assume that the individual's supply decisions are made on a purely economic basis; there are no net advantages or disadvantages to the production of medical services. If supply is inelastic, x_1 and x_2 are given, and p is determined by demand conditions. In the case where medical services are produced perfectly elastically at price p, $x_1 + px_2$ is given to the individual and to society, p is determined by technological considerations, and the actual values of x_1 and x_2 are determined by demand conditions.

In the general imperfectly elastic supply case, the supply functions $x_1(p)$ and $x_2(p)$ are determined so as to maximize $x_1 + px_2$ subject to a transformation constraint. Since p is the marginal rate of transformation,

$$\frac{dx_1}{dp} + p\,\frac{dx_2}{dp} = 0 \ . \tag{4}$$

Once the state s has occurred, the individual maximizes $U(x_{1s}, x_{2s}, s)$ subject to the constraint

$$x_{1s} + qx_{2s} = y \ . \tag{5}$$

(Strictly speaking, the budget constraint is a weak inequality, but in this case it is assumed to be binding). The optimality conditions then are (5) and the equations

$$\frac{\partial U}{\partial x_{1s}} = \lambda_s, \ \frac{\partial U}{\partial x_{2s}} = \lambda_s q \ , \tag{6}$$

where λ_s is the marginal utility of income in state s.

The optimization defines, for each s, the demand functions $x_{1s}(q, y)$ and $x_{2s}(q, y)$. For given r, q and y are in turn functions of p, x_1, and x_2. The aggregate demands, $E(x_{is})$ (i = 1, 2) are therefore also functions of these variables. The two equations (1) are not independent; if the equation for i = 2 is multiplied by q and added to that for i = 1,

$$E(x_{1s} + qx_{2s}) = x_1 + qx_2 \ ;$$

but from (3) and (5), this reduces to the tautology, E(y) = y (recall that y is not dependent on s). Hence the equilibrium values of p, x_1, and x_2 are defined by one of the equations (1) together with the supply conditions. There is thus an equilibrium allocation of resources between medical care and other goods for each value of the coinsurance rate r, and I wish to evaluate these alternative equilibria. Ideally, I would like to optimize on r; as a minimum, I would like to determine whether an increase or a decrease in the coinsurance rate would increase welfare.[1]

The criterion of welfare is taken to be the expected utility of the representative individual. This respects his attitude to risk aversion; it also has the implication of respecting his tradeoff between medical care and other goods in any given state of health, a point that might be more arguable but will not be challenged here. Let W be the individual's welfare:

$$W = E[U(x_{1s}, x_{2s}, s)] \ . \tag{7}$$

[1]Of course, this is within the context of insurance schemes that are purely linear in medical costs. Indeed, actual insurance plans, with their deductibles followed by coinsurance, are nonlinear and need closer investigation.

The equilibrium magnitudes of the system are functions of r. In particular, for each s, x_{1s} and x_{2s} are functions of q and y, which in turn are determined, for fixed r, by p, x_1, and x_2. Hence, W is a function of r. I shall examine the effects of marginal changes in W; the spirit of this analysis is therefore very much that of Lesourne (1972, Ch. 2, Sec. I).

III. THE GENERAL FORMULA FOR WELFARE EFFECTS: FIRST FORM

First, differentiate the budget equation (5) with respect to r.

$$\frac{dx_{1s}}{dr} + q\,\frac{dx_{2s}}{dr} = \frac{dy}{dr} - x_{2s}\,\frac{dq}{dr}\,.$$

But from (3),

$$\frac{dy}{dr} = \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + x_2\,\frac{dq}{dr}\,,$$

so that

$$\frac{dx_{1s}}{dr} + q\,\frac{dx_{2s}}{dr} = \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + (x_2 - x_{2s})\,\frac{dq}{dr}\,. \qquad (8)$$

Now differentiate the welfare criterion (7) with respect to r.

$$\frac{dW}{dr} = E\left(\frac{\partial U}{\partial x_{1s}}\,\frac{dx_{1s}}{dr} + \frac{\partial U}{\partial x_{2s}}\,\frac{dx_{2s}}{dr}\right).$$

First substitute from (6) and then from (8):

$$\begin{aligned}\frac{dW}{dr} &= E\left[\lambda_s\left(\frac{dx_{1s}}{dr} + q\,\frac{dx_{2s}}{dr}\right)\right]\\ &= E\left[\lambda_s\left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + (x_2 - x_{2s})\,\frac{dq}{dr}\right)\right].\end{aligned}$$

Note, however, that the magnitudes

$$\frac{dx_1}{dr}\,,\ \frac{dx_2}{dr}\,,\ \text{and}\ \frac{dq}{dr}$$

are independent of the state of health and hence are not random variables. They can therefore be factored out of any expectation. The marginal effect of coinsurance rate on welfare therefore takes the form

$$\frac{dW}{dr} = \left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\right) E(\lambda_s) - E[\lambda_s(x_{2s} - x_2)]\,\frac{dq}{dr}\,. \tag{9}$$

The first term of (9) represents the welfare gain within each state s due to an increase in the coinsurance rate, which is a decrease in the subsidy to the consumption of medical services. The second term is the distinctive element that measures the welfare loss because of increased risk-bearing. With regard to the second term, note that from (1), x_2 is the mean value of x_{2s}; hence, by the usual definition of a covariance,

$$E[\lambda_s(x_{2s} - x_2)] = \sigma_{\lambda_s x_{2s}}\,, \tag{10}$$

the covariance of medical services used with the marginal utility of income. I shall return to this term in Section V.

At the moment, I use standard methods of second-best analysis (see, for example, the discussion in Lesourne, 1972, cited above) to restate the first factor in the first term of (9). This is the conventional measure of marginal welfare effect, and I give it the symbol

$$W_o = \frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\,. \tag{11}$$

Because of the presence of uncertainty and the fact that markets are therefore not perfect (in the sense that the full set of contingent markets is not available), there are two somewhat different expressions that can be found for W_o, the first of which is more useful econometrically and the second of which is more useful for theoretical analysis.

Since x_{2s} is a function of q and y, and with the aid of the definition of income (3),

$$\frac{dx_{2s}}{dr} = \frac{\partial x_{2s}}{\partial q}\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\frac{dy}{dr}$$

$$= \frac{\partial x_{2s}}{\partial q}\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + x_2\,\frac{dq}{dr}\right)$$

$$= \left(\frac{\partial x_{2s}}{\partial q} + x_2\,\frac{\partial x_{2s}}{\partial y}\right)\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\right) . \tag{12}$$

Take expectations of both sides of (12). From (1),

$$E\left(\frac{dx_{2s}}{dr}\right) = \frac{dE(x_{2s})}{dr} = \frac{dx_2}{dr} ,$$

and therefore,

$$\frac{dx_2}{dr} = \frac{dq}{dr}\left[E\left(\frac{\partial x_{2s}}{\partial q}\right) + x_2\,E\left(\frac{\partial x_{2s}}{\partial y}\right)\right] + \left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\right)E\left(\frac{\partial x_{2s}}{\partial y}\right) .$$

Now, from (4) and (11),

$$\begin{aligned} W_o &= \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} = \frac{dx_1}{dr} + p\,\frac{dx_2}{dr} + (q - p)\,\frac{dx_2}{dr} \\ &= (q - p)\,\frac{dx_2}{dr} \\ &= (q - p)\,\frac{dq}{dr}\left[E\left(\frac{\partial x_{2s}}{\partial q}\right) + x_2\,E\left(\frac{\partial x_{2s}}{\partial y}\right)\right] \qquad (13) \\ &\quad + (q - p)\,E\left(\frac{\partial x_{2s}}{\partial y}\right)W_o . \end{aligned}$$

The expression

$$\bar{S}_{22} = E\left(\frac{\partial x_{2s}}{\partial q}\right) + x_2\,E\left(\frac{\partial x_{2s}}{\partial y}\right) \tag{14}$$

resembles a Slutsky compensated derivative but, in fact, is not one nor is it the expectation of one. However, it can also be written

$$\bar{S}_{22} = \frac{\partial E(x_{2s})}{\partial q} + E(x_{2s}) \frac{\partial E(x_{2s})}{\partial y} ;$$

if a demand curve for medical services is fitted to time series in the usual way, the dependent variable is $E(x_{2s})$, and therefore all the terms in the expression can be calculated from the econometric analysis.

Solve in (13) for W_o, using the abbreviation (14).

$$W_o = \frac{(q - p)\, \bar{S}_{22}}{1 + (p - q)\, E\left(\frac{\partial x_{2s}}{\partial y}\right)} \frac{dq}{dr} . \qquad (15)$$

When there is some insurance, $q < p$; since medical services are a normal good, the denominator is certainly positive, and the sign of W_o is opposite to that of $\bar{S}_{22}$ (provided $dq/dr > 0$).

An alternative expression for $\bar{S}_{22}$ will strongly suggest that it must be negative. Let S_{22s} be the compensated effect of a change in the price of medical services for a given state s--that is, the derivative of x_{2s} with respect to q when the consumer remains on an indifference curve for that state.

$$S_{22s} = \frac{\partial x_{2s}}{\partial q} + x_{2s} \frac{\partial x_{2s}}{\partial y} .$$

Then

$$\frac{\partial x_{2s}}{\partial q} + x_2 \frac{\partial x_{2s}}{\partial y} = S_{22s} - (x_{2s} - x_2) \frac{\partial x_{2s}}{\partial y} .$$

Taking expectations,

$$\bar{S}_{22} = E(S_{22s}) - E\left[(x_{2s} - x_2) \frac{\partial x_{2s}}{\partial y}\right] .$$

Of course, for each s, $S_{22s} < 0$, so the first term is negative. The second term is the covariance between medical services and the marginal propensity to consume them (remember this is the covariance across states of health for a given individual). This term can also, and perhaps more illuminatingly, be rewritten as follows: Since,

$$(x_{2s} - x_2)\,\frac{\partial x_{2s}}{\partial y} = 1/2\,\frac{\partial (x_{2s} - x_2)^2}{\partial y}\ ,$$

$$E\left[(x_{2s} - x_2)\,\frac{\partial x_{2s}}{\partial y}\right] = 1/2\ E\left[\frac{\partial (x_{2s} - x_2)^2}{\partial y}\right] = 1/2\,\frac{\partial E[(x_{2s} - x_2)^2]}{\partial y}$$

$$= 1/2\,\frac{\partial \sigma^2_{x_{2s}}}{\partial y}\ .$$

$$\overline{S}_{22} = E(S_{22s}) - 1/2\,\frac{\partial \sigma^2_{x_{2s}}}{\partial y}\ . \qquad (16)$$

The second term is rather unexpected; it represents the effect of income on the variance of health expenditure. It seems reasonable to assume that a higher income permits higher medical expenditures in more serious illnesses; hence one would expect that the variance of medical expenditures would increase with income. Therefore it is to be presumed *a fortiori* that $\overline{S}_{22}$ is negative and that $W_o > 0$ when $q > p$. Phelps has taken the absolute values of the residuals from a regression of physician visits on a number of variables including income (see Newhouse and Phelps, 1973) and shown that the correlation with income is slightly negative, in contrast to this argument. However, the effect is small compared with the first term in (16), so that the negativity of $\overline{S}_{22}$ is not in question.

<u>REMARK 1</u>. Clearly, when $q = p$ (no insurance at all), then W_o vanishes completely.

<u>REMARK 2</u>. If $dq/dr = 0$, then again $W_o = 0$. Then by (11) the first term in (9) is 0 and so is the second term, so that $dW/dr = 0$. To bring this out more clearly, note that the specific definition of r played no role in the analysis; any parameter of the insurance contract would have yielded the same formulas. In particular, r might have been replaced by q everywhere. That is, one could imagine an

insurance system in which the government chose the buyer's price, rather than a coinsurance rate, and then let the forces of the market determine seller's price and from that the needed lump-sum taxes. The analysis would have proceeded along the same lines, except that dq/dr would have been replaced by the number 1. If

$$W_o' = \frac{dx_1}{dq} + q \frac{dx_2}{dq} \quad , \tag{17}$$

then the analogue of (15) is

$$W_o' = \frac{(q - p)\ \bar{S}_{22}}{V} \text{ , where } V = 1 + (p - q)\ E\left(\frac{\partial x_{2s}}{\partial y}\right) \text{ ,} \tag{18}$$

and (15) itself becomes

$$W_o = W_o' \frac{dq}{dr} \text{ .} \tag{19}$$

From (9), (10), (11), (18), and (19), one can write

THEOREM 1. A general formula for the marginal effect of an increase in the coinsurance rate on expected welfare is

$$\frac{dW}{dr} = [W_o'\ E(\lambda_s) - \sigma_{\lambda_s x_{2s}}]\ \frac{dq}{dr} \text{ ,}$$

where

$$W_o' = \frac{(q - p)\ \bar{S}_{22}}{V} \text{ , } V = 1 + (p - q)\ E\left(\frac{\partial x_{2s}}{\partial y}\right) \text{,}$$

and $\bar{S}_{22}$ is given by either of the following formulas:

$$\bar{S}_{22} = \frac{\partial E(x_{2s})}{\partial q} + E(x_{2s})\ \frac{\partial E(x_{2s})}{\partial y}$$

$$= E(S_{22s}) - 1/2\ \frac{\partial \sigma^2_{x_{2s}}}{\partial y} \text{ ,}$$

and S_{22s} is the compensated effect of a change in the price of medical services within a given state. I shall argue in Section V that under plausible assumptions the covariance between marginal utility of income and medical services is positive.

It should be noted that the welfare effect in Theorem 1 is measured in utility terms and therefore in arbitrary units. Usually welfare losses are measured in some convenient numeraire. In this case, there is no completely obvious numeraire. Goods other than medical services appear to be the obvious choice, but under uncertainty this is not a well-defined commodity; one has to distinguish among other goods in different states of health. The simplest numeraire appears to be the composite good consisting of one unit of "other goods" in every state of health. The marginal utility of this composite is $E(\lambda_s)$, and therefor the welfare loss measured in other goods is obtained by dividing dW/dr by $E(\lambda s)$.

IV. THE GENERAL FORMULA FOR WELFARE EFFECTS: TAKING EXPLICIT ACCOUNT OF SUPPLY FACTORS

The formulas in Theorem 1 made no use of the fact that r was to be interpreted as the coinsurance rate; they are valid for any shift in the insurance scheme or indeed in any other parameter. The specific effect of coinsurance rates is confined to the expression dq/dr, the effect of the coinsurance rate on the buyer's price of medical services. The evaluation requires supply considerations.

In the case of perfectly elastic supply, p, the seller's price, is given by the technology. Since $q = rp$,

$$\frac{dq}{dr} = p \text{ when supply is perfectly elastic.} \qquad (20)$$

Then the expressions in Theorem 1 can be evaluated from demand considerations alone.

THEOREM 2. When supply of medical services is perfectly elastic, production of a unit of medical services requires giving up p units of other goods. Then the marginal welfare effect of an increase in the coinsurance rates is

$$\frac{dW}{dr} = [W_o' \, E(\lambda_s) - \sigma_{\lambda_s x_{2s}}] \, p \, ,$$

where W_o' is defined in the statement of Theorem 1.

The welfare evaluations of medical insurance that have been made (for example, Feldstein, 1973; Pauly, 1968) have assumed perfect elasticity.

The imperfectly elastic case, at least in the short run, is much more realistic. Even in the long run, the production of both physicians and hospital services occurs under such special circumstances that perfect elasticity cannot be taken for granted.

Assume that the supplies of the two types of goods are functions of $p = q/r$; in particular, x_1 and x_2 are functions of p. Let

$$x_i' = \frac{dx_i}{dp} \quad (i = 1, 2) \ .$$

Differentiate the equation (1)

$$E(x_{2s}) = x_2(p) \ ,$$

with respect to r, to solve for dq/dr.

$$E\left(\frac{\partial x_{2s}}{\partial q}\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\frac{dy}{dr}\right) = x_2' \frac{dp}{dr} \ . \tag{21}$$

From (3)

$$\frac{dy}{dr} = (x_1' + qx_2') \frac{dp}{dr} + x_2 \frac{dq}{dr} \ .$$

Any shift in supplies induced by a change in p must lie on the transformation surface; by (4),

$$x_1' + px_2' = 0 \ ,$$

hence, as before,

$$x_1' + qx_2' = (q - p)\, x_2' \ .$$

From (21)

$$E\left(\frac{\partial x_{2s}}{\partial q} + x_2 \frac{\partial x_{2s}}{\partial y}\right)\frac{dq}{dr} = x_2' \left[1 + (p - q)\, E\left(\frac{\partial x_{2s}}{\partial y}\right)\right] \frac{dp}{dr} \ ,$$

or, with the notation defined in the statement of Theorem 1,

$$\bar{S}_{22} \frac{dq}{dr} = x_2' \, V \frac{dp}{dr}. \tag{22}$$

From the identity, $p = q/r$,

$$\frac{dp}{dr} = \frac{1}{r}\frac{dq}{dr} - \frac{q}{r^2} .$$

Substitute in (22), and solve for dq/dr.

$$\frac{dq}{dr} = \frac{pVx_2'}{x_2' \, V - r\,\bar{S}_{22}} . \tag{23}$$

From (23) the following observations are immediate:

(1) since $\bar{S}_{22} < 0$, $dq/dr < p$, if x_2' is finite;

(2) if x_2' is infinite (perfect elasticity), $dq/dr = p$;

(3) if $x_2' = 0$ (perfect inelasticity), $dq/dr = 0$;

(4) as r approaches 0, dq/dr approaches p (this assumes that there is satiation in demand and some elasticity of supply, so that there is a finite equilibrium value of p).

Substituting (23) into Theorem 1, after some simplification, yields

THEOREM 3. Let $x_2(p)$ be the supply function of medical services. Then,

$$\frac{dW}{dr} = \frac{[(q - p)\,\bar{S}_{22}\,E(\lambda_s) - V\,\sigma_{\lambda_s x_{2s}}]\,px'_2}{Vx_2' - r\,\bar{S}_{22}}$$

where V and $\bar{S}_{22}$ are defined as in the statement of Theorem 1. The evaluation of this expression depends on econometric estimation of the demand and supply curves and on the evaluation of the covariance term.

V. COVARIANCE BETWEEN MARGINAL UTILITY OF INCOME AND MEDICAL SERVICES

To show that this covariance is positive, it suffices to indicate that both are increasing functions of the state of health. This presupposes that the states of health are measured in order of increasing severity of illness (poor health has a high index number). The desired result is derived from the following assumptions:

A.1. For fixed levels of medical services and other goods, the marginal rate of substitution of other goods for medical services is an increasing function of the state of health.

A.2. For fixed levels of medical services and other goods, the marginal utility of other goods does not decrease with the state of health.

A.3. For a fixed state of health, the utility is jointly concave in medical services and other goods and twice continuously differentiable.

A.4. For a fixed state of health and a fixed level of medical services, the marginal rate of substitution of other goods for medical services increases with an increase in the amount of other goods.

A.1. amounts to saying that the states of health are ordered in such a way as to make it true; the assumption is not tautological because it does assert that if the marginal rate of substitution increases from one state to another for one pair of values of medical services and other goods, it does so for all.

The meaning of A.2 is that if an individual is initially given a fixed level of other goods and of medical services for all states of health, he would prefer to switch, if at all, to having the level of other goods rise with illness, if the switch can be made on an actuarially fair basis and if the level of medical services in any state is not subject to change. As Joseph Newhouse has pointed out to me, A.2 is not expected to hold for all states of health (see also Arrow, 1973,

p. 3). Certainly, some states of health sharply reduce the value of other goods to the patient; he is too ill to enjoy the consumption and *ex ante* would have preferred to have shifted consumption of other goods to states of better health. In many states of ill health, however, other goods (such as domestic servants and other forms of service, or travel to less demanding climates) may be valued very highly. The result contained in Theorem 3 remains valid if A.4 holds only on the average or even if the marginal utility of other goods falls, but not too rapidly, as the state of health deteriorates.

A.3 is a usual statement of risk aversion. It means that given any two possible pairs (x_1^o, x_2^o) and (x_1^1, x_2^1) of other goods and medical services, the individual would prefer their average to an even chance on getting one or the other.

A.4 is self-explanatory.

I first derive expressions for the rates of change of medical services and of the marginal utility of income with respect to state of health and then show that, under the above assumptions, both are positive.

Differentiate the optimality conditions (6) and the budget equation (5) with respect to s, the state of health.

$$U_{11} \frac{dx_{1s}}{ds} + U_{12} \frac{dx_{2s}}{ds} + \left(- \frac{d\lambda_s}{ds}\right) = - U_{1s} ,$$

$$U_{21} \frac{dx_{1s}}{ds} + U_{22} \frac{dx_{2s}}{ds} + q \left(- \frac{d\lambda_s}{ds}\right) = - U_{2s} ,$$

$$\frac{dx_{1s}}{ds} + q \frac{dx_{2s}}{ds} = 0 ,$$

where

$$U_{ij} = \frac{\partial^2 U}{\partial x_{is} \partial x_{js}} \ (i, j = 1, 2), \ U_{is} = \frac{\partial^2 U}{\partial x_{is} \partial s} \ (i = 1, 2) .$$

We can treat this system in the usual way as linear in the derivatives, dx_{1s}/ds, dx_{2s}/ds, and $-\,d\lambda_s/ds$. From the second-order conditions for a constrained optimum, the determinant, D, of the above system must be positive. Straightforward use of Cramer's rule yields

$$\frac{dx_{2s}}{ds} = \frac{U_{2s} - q\,U_{1s}}{D}\,, \tag{24}$$

$$\frac{d\lambda_s}{ds} = \frac{U_{1s}\,(qU_{21} - U_{22}) + U_{2s}\,(U_{12} - qU_{11})}{D}\,. \tag{25}$$

Since $q = U_2/U_1$, the numerator of (24) can be written

$$U_2\left(\frac{U_{2s}}{U_2} - \frac{U_{1s}}{U_1}\right) = U_2\,\frac{\partial \log\,(U_2/U_1)}{\partial s}\,.$$

Since $U_2 > 0$ and, from A.1, U_2/U_1 is increasing in s,

$$\frac{dx_{2s}}{ds} > 0\,.$$

Equivalently,

$$U_{2s} > qU_{1s}\,.$$

From A.4,

$$U_{12} - qU_{11} = U_2\left(\frac{U_{21}}{U_2} - \frac{U_{11}}{U_1}\right) = U_2\,\frac{\partial \log\,(U_2/U_1)}{\partial x_{1s}} > 0\,.$$

From the last two relations, the numerator of (25) satisfies the inequality

$$U_{1s}\,(qU_{21} - U_{22}) + U_{2s}\,(U_{12} - qU_{11}) > U_{1s}\,(qU_{21} - U_{22})$$

$$+ qU_{1s} (U_{12} - qU_{11})$$

$$= - U_{1s} (U_{11} q^2 - 2U_{12} q + U_{22}) \geq 0 ,$$

since $U_{1s} \geq 0$ by A.2, and

$$U_{11} q^2 - 2U_{12} q + U_{22} \leq 0 ,$$

by A.3.

<u>THEOREM 4</u>. Under assumptions A.1-A.4, the marginal utility of income is positively correlated with medical expenditures. The covariance constitutes an offsetting risk adjustment to the marginal welfare change with respect to increase in coinsurance.

In particular Theorem 4 establishes that some insurance is better than no insurance. From Theorem 1 and the fact that $V = 1$ when $r = 1$--that is, $q = p$,

$$\frac{dW}{dr} = - \sigma_{\lambda_s x_{2s}} \frac{dq}{dr} < 0 ,$$

provided $dq/dr > 0$.

It appears that nothing further can be estimated on a theoretical basis, except for the special case of inelastic supply discussed in the next section. It is not even excluded, so far as I can see, that complete insurance be optimal, although it is unlikely. In that case, $q = 0$; hence, the budget constraint tells us that $x_{1s} = y$, and x_{2s} is determined by the condition

$$\frac{\partial U_s}{\partial x_{2s}} = 0 .$$

Since medical care is always costly in discomfort and time, we can suppose the demand will be satiable. The solution to the last equation, when $x_{1s} = y$, will be denoted by x^o_{2s}. When $q = 0$, it is easy to calculate that $D = - U_{22}$. Hence (24) and (25) become

$$\frac{dx^o_{2s}}{ds} = - \frac{U_{2s}}{U_{22}} , \frac{d\lambda_s}{ds} = U_{1s} - \frac{U_{2s}\ U_{12}}{U_{22}} = U_{1s} + U_{12}\frac{dx^o_{2s}}{ds} .$$

The consumption of free medical services is certainly increasing with the state of health (measured to increase with increasing illness); indeed, since the state of health has so far appeared only ordinally, it is reasonable to identify x^o_{2s} with s, so that $dx^o_{2s}/ds = 1$. Hence,

$$\frac{d\lambda_s}{ds} = U_{1s} + U_{12} , \sigma_{\lambda_s x_{2s}} = \sigma_{\lambda_s s} .$$

The relation between marginal utility of income and state of health when medical care is free depends on the cross-effects of state of health and of medical services on the marginal utility of other goods. It can be shown (see Appendix A) that the covariance in question equals the variance of free medical services multiplied by an average value of the derivative $d\lambda_s/ds = U_{1s} + U_{12}$; in symbols,

$\sigma_{\lambda_s s} = U\ \sigma^2_s$, where U is a weighted average of the values of $U_{1s} + U_{12}$ for varying s.

The risk-aversion term may not vanish even for zero coinsurance. (Remember this is a term in the *marginal* welfare effect; the risk-aversion welfare gain is of the first order in the coinsurance rate.) It is therefore conceivable that it outweighs the allocation term. In general the values of U_{1s} and U_{12} should be small, so perfect insurance should not be optimal.

The calculation of $\sigma_{\lambda_s x_{2s}}$ can be made only by assuming specific forms for the utility function and the distribution of states of health. A specific example is developed in Appendix B.

VI. THE CASE OF PERFECTLY INELASTIC SUPPLY

In welfare economics we are accustomed to the argument that when supply is totally inelastic, changes in prices have no welfare effects. The argument may need to be reexamined here, because of the presence of uncertainty in demand and the absence of contingent markets; but the conclusion remains valid. It does not seem to have been adequately remarked in the literature on health insurance that *if the supply of medical care is perfectly inelastic, then there is no welfare effect at all from a change in the coinsurance rate*. There is, however, a rise in the price of medical services paid to the supplier; in a multi-person world this amounts to a redistribution of income to the suppliers of medical services.

This conclusion follows immediately by setting $x_2' = 0$ in Theorem 3. Note that if x_1 and x_2 are both given, either of the equations (1) has only a single unknown, q; r, the coinsurance rate, does not enter, and income, y, is determined by q, from (3). The equilibrium buyer's price and income are the same for all values of r; in particular, the demands for medical services and for other goods in each state s is the same for all r, and therefore expected utility is independent of r.

The only variable that does change with changing r is p, since $p = q/r$. That is, the price of medical services rises as coinsurance rates fall. The pre-tax income of society is increasingly directed to medical services. To the extent that taxes to pay for medical services do not fall on medical income, there is a transfer of income to the suppliers of medical services. To illustrate, suppose that the cost of medical insurance is paid for by a proportional income tax at a rate t.

$$t = \frac{(p - q)\ x_2}{x_1 + px_2} = \frac{(1 - r)\ qx_2}{rx_1 + qx_2} .$$

Then, the ratio of post-tax nonmedical incomes to their level with no insurance is

$$\frac{(1 - t)\, x_1}{x_1} = 1 - t = \frac{r(x_1 + qx_2)}{rx_1 + qx_2} \quad ,$$

which decreases from 1 toward 0 as r decreases from 1 to 0. Correspondingly, the ratio of post-tax medical incomes to their no-insurance level is

$$\frac{(1 - t)\, px_2}{qx_2} = \frac{1 - t}{r} = \frac{x_1 + qx_2}{rx_1 + qx_2} \; ,$$

which rises from 1 as r decreases.

Appendix A

COVARIANCE OF MARGINAL UTILITY OF INCOME AND HEALTH

THEOREM. If X is a random variable and f(X) is a **function**, then $\sigma_{f(X)X} = u\,\sigma_X^2$, where u is a weighted average of f'(X).

Proof. Let g(x) be the density of X, a and b the limits of the range of X, G(x) the cumulative distribution of X, and

$$H(x) = \int_a^x y\, g(y)\, dy\ .$$

$$\sigma_{f(X)X} = E[f(X)X] - E(X)\, E[f(X)]\ .$$

Integrating by parts,

$$E[f(X)X] = \int_a^b f(x)\, x\, g(x)\, dx = f(b)\, H(b) - f(a)\, H(a)$$

$$- \int_a^b f'(x)\, H(x)\, dx\ .$$

$$E[f(X)] = \int_a^b f(x)\, g(x)\, dx = f(b)\, G(b) - f(a)\, G(a)$$

$$- \int_a^b f'(x)\, G(x)\, dx.$$

By definition,

$$H(b) = E(X),\ H(a) = 0,\ G(b) = 1,\ G(a) = 0\ .$$

Hence,

$$\sigma_{f(X)X} = \int_a^b f'(x)\ [E(X)\ G(x) - H(x)]\ dx\ .$$

Let

$$W(X) = E(X)\ G(x) - H(x)\ .$$

Then,

$$\sigma_{f(X)X} = \int_a^b f'(x)\ W(x)\ dx\ .$$

This holds for any function f(X). In particular, let f(X) = X, so that f' = 1.

$$\sigma_X^2 = \int_a^b W(x)\ dx\ ,$$

so that

$$\sigma_{f(X)X} \ /\ \sigma_X^2 = \int_a^b f'(x)\ w(x)\ dx,$$

where

$$w(x) = W(x) \Big/ \int_a^b W(x)\ dx\ .$$

By construction,

$$\int_a^b w(x)\ dx = 1\ .$$

To show that $u = \sigma_{f(X)X} / \sigma_X^2$ is a weighted average of $f'(X)$, it suffices to show that $w(x)$ is nonnegative, or, equivalently, that $W(x)$ is nonnegative.

Differentiating,

$$W'(x) = E(X)\ g(x) - x\ g(x) = g(x)\ [E(X) - x] .$$

Hence, $W(x)$ is increasing for $x < E(X)$ and decreasing for larger values of x. It has a maximum at $x = E(X)$ and minima at the extremes, $x = a$ and $x = b$. But $W(a) = 0$, $W(b) = E(X) - H(b) = E(X) - E(X) = 0$, so $W(x) > 0$ for all x, $a < x < b$.

In the text, X is interpreted as the state of health s (as measured by the consumption of medical services, x_{2s}^o, when free) and $f(X)$ as the marginal utility of income, λ_s, with the derivative $d\lambda_s/ds = U_{1s} + U_{12}$.

Appendix B

COVARIANCE OF MARGINAL UTILITY OF INCOME AND MEDICAL SERVICES FOR A SPECIFIC UTILITY FUNCTION AND DISTRIBUTION OF STATES OF HEALTH

I seek here to illustrate how expressions might be found for the covariance term of Section V provided assumptions are made about the nature of the utility function and the distribution of medical services for a given coinsurance rate.

Assume that

$$U(x_{1s}, x_{2s}, s) = -(1/c)\, e^{-cx_{1s}} + U_2\,(x_{2s}, s)\ .$$

That is, I assume (a) that utility is additive in other goods and in medical services, and (b) the utility function for other goods has constant absolute risk aversion (this assumption is made by Feldstein, 1973). Assume further that the distribution of medical services for a given income and a given coinsurance rate is described by a gamma-distribution (see Friedman, 1971),

$$\frac{a^b}{\Gamma(b)}\, e^{-ax_{2s}}\, (x_{2s})^{b-1}\ .$$

Then, for any numbers m, n,

$$E(e^{-mX_{2s}}\, X_{2s}^n) = \frac{a^b}{\Gamma(b)} \int_0^{+\infty} e^{-mx_{2s}}\, x_{2s}^n\, e^{-ax_{2s}}\, x_{2s}^{b-1}\, dx_{2s}$$

$$= \frac{a^b}{\Gamma(b)} \int_0^{+\infty} e^{-(a+m)x_{2s}}\, x_{2s}^{n+b-1}\, dx_{2s}\ .$$

Let $y = (a+m)\, x_{2s}$.

$$E(e^{-mX_{2s}} X_{2s}^n) = \frac{a^b}{\Gamma(b)} \frac{1}{(a+m)^{n+b}} \int_o^{+\infty} e^{-y} y^{n+b-1} dy \tag{B.1}$$

$$= \frac{a^b \Gamma(n + b)}{\Gamma(b) (a + m)^{n+b}} .$$

If m = 0, n = 1, we have

$$E(X_{2s}) = \frac{\Gamma(b + 1) a^b}{\Gamma(b) a^{b+1}} = \frac{b}{a} .$$

Since

$$\lambda_s = \frac{\partial U}{\partial x_{1s}} = e^{-cx_{1s}} = e^{-c(y - qx_{2s})} = e^{-cy} e^{cqx_{2s}} ,$$

with the aid of the budget constraint,

$$E(\lambda_s) = e^{-cy} E(e^{cqx_{2s}}) = \frac{a^b \Gamma(b) e^{-cy}}{\Gamma(b)(a - cq)^b} = \frac{a^b e^{-cy}}{(a - cq)^b}$$

from (B.1), with m = -cq and n = 0.

Then,

$$\sigma_{\lambda_s x_{2s}} = E(\lambda_s X_{2s}) - E(\lambda_s) E(X_{2s})$$

$$= E(e^{-cX_{1s}} X_{2s}) - \frac{ba^{b-1} e^{-cy}}{(a - cq)^b}$$

$$= e^{-cy} E(e^{cqX_{2s}} X_{2s}) - \frac{ba^{b-1} e^{-cy}}{(a - cq)^b}$$

$$= e^{-cy}\left[\frac{a^b\Gamma(b + 1)}{\Gamma(b)(a - cq)^{b+1}} - \frac{ba^{b-1}}{(a - cq)^b}\right]$$

$$= \frac{e^{-cy}\, ba^{b-1}\, cq}{(a - cq)^{b+1}} ,$$

where use is made of (B.1) with $m = -cq$ and $n = 1$.

This calculation is designed to show merely that manageable formulas are not impossible, even though at the cost of strong assumptions. The parameters a and b of the distribution of medical services demanded are in principle observable. The absolute risk aversion cannot be inferred from data on the demand for medical services but is at least inferrable for observed behavior in the presence of uncertainty--for example, the choice of stock portfolios. The assumption of constant absolute risk aversion is uncomfortable; it implies, for example, that the demand for risky assets does not increase with wealth. However, alternative assumptions, such as constant relative risk aversion, do not lead to simple formulas, though in any case they always lead to expressions that can be evaluated numerically.

REFERENCES

Arrow, K. J., *Optimal Insurance and Generalized Deductibles*, R-1108-OEO, The Rand Corporation, February 1973. To appear in *Skandinavsk Aktuarietidskrift*, 1974.

Feldstein, M., "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, 81:251-280, 1972.

Friedman, B. S., *A Study of Uncertainty and Health Insurance*, Ph.D. dissertation, Massachusetts Institute of Technology, 1971.

Lesourne, J., *Le Calcul Économique*, 2e édition, Dunod, Paris, 1972.

Newhouse, J. P. and C. E. Phelps, *Price and Income Elasticities for Medical Care Services*, R-1197-NC, The Rand Corporation (forthcoming).

Pauly, M., "The Economics of Moral Hazard," *American Economic Review*, 53:531-537, 1968.